

Multi-Sensory, Multi-Modal Concepts for Information Understanding

David L. Hall, Ph.D.

School of Information Sciences and Technology
The Pennsylvania State University
268 Homan Avenue
State College, PA 16801
USA

dhall@ist.psu.edu

Timothy S. Shaw

Applied Research Laboratory
The Pennsylvania State University
University Park
Pennsylvania
USA

tshaw@psu.edu

ABSTRACT

In recent years, extensive research has focused on the development of techniques for multi-sensor data fusion systems. These fusion systems process data from multiple sensors to develop improved estimates of the position, velocity, attributes, and identity of entities such as targets or entities of interest. Typically, the fused data are displayed on a geographical information display (e.g., data are overlaid on a map with terrain features, political information, and the data are shown as icons representing the observed entities). Analysts interpret the data to develop an assessment of an evolving situation or threat. Despite significant improvements in computer displays, haptic interface devices, and new three-dimensional full-immersion display capabilities, the data fusion displays have seen little changes to take advantage of a human's ability to access data. This paper describes two concepts for improved understanding of data; (1) the utilization of multiple human senses to interact with and interpret data, and (2) the dual use of language and vision to improve information understanding.

INTRODUCTION

Historically, information displays for display and understanding of data fusion products have focused on the use of vision. These displays typically involve geographical information displays involving a map and map overlays. The evolution of such displays has focused on increasing resolution, multi-layer information displays, utilization of icons, and interaction mechanisms using pull-down windows and related techniques. We argue that, while vision is a powerful human sense, such interfaces ignore the other human senses of hearing, touch, smell and taste. True virtual reality would exploit all of these senses to make displayed data seem more “real”, and to take advantages of individual preferences for information access and improved understanding via multi-mode interaction. Research conducted by our group at Penn State University has utilized multiple human senses (including vision, sound and touch) to improve the analysis and understanding of data. New techniques have been developed including; (1) use of sound sonification to interpret data uncertainty, (2) use of novel 3D visual devices paradigms, such as height to represent temporal phenomena, (3) use of transparency to represent uncertainty, and (4) haptic devices to interpret data. The utilization of such methods provides the opportunity for improved understanding of complex data such as multi-spectral, multi-sensor environmental data.

The second area of exploration has involved the combined utilization of human visual intelligence and language intelligence to understand data. Humans exhibit intelligence in two broad arenas; vision and language skills. Standard data fusion displays rely only on human visual intelligence to interpret data. Examples of visual intelligence include pattern recognition, understanding of proximity and motion, and utilization of our knowledge of the physical world. Humans have the ability to apply a visual intelligence

Paper presented at the RTO IST Workshop on “Massive Military Data Fusion and Visualisation: Users Talk with Developers”, held in Halden, Norway, 10-13 September 2002, and published in RTO-MP-105.

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 00 APR 2004		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Multi-Sensory, Multi-Modal Concepts for Information Understanding				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) School of Information Sciences and Technology The Pennsylvania State University 268 Homan Avenue State College, PA 16801 USA; Applied Research Laboratory The Pennsylvania State University University Park Pennsylvania USA				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM001665, RTO-MP-105 Massive Military Data Fusion and Visualization: Users Talk with Developers., The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 26	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

to interpret information. While this is a powerful cognitive skill, there are cognitive biases based on vision. Examples include our lack of ability to “un-see” a pattern once it has been recognized, identification of patterns in random data (recognition of patterns even when patterns are not meaningful), and an implicit assumption that proximity is related to causal relationships. We have developed some concepts to explicitly use both vision and language capability for improved information understanding. These concepts leverage some newly developed methods for transforming image information into text (viz., descriptions about the image content). Use of such transformations allows a user to access and use both visual intelligence and natural language reasoning ability.

AN EXPERIMENT WITH NASA DATA

Some of the concepts described above were explored under a grant from the National Aeronautics and Space Administration (NASA [1]). The objective of the NASA funded research into information fusion was to combine remote sensor data with other information sources for improved data interpretation and understanding. We focused on automated contextual interpretation of multi-source information integration and fusion supporting human analysis [2].

For this project, information fusion techniques include the correlation and conditioning of data products, both geo-spatially and temporally, and fusion and interpretation of data using a hybrid reasoning approach. The following topics were pursued:

- 1) Fusion of non-commensurate multi-source sensor data and information, including images and textual reports.
- 2) Performing automated contextual interpretation of earth science data to assist human analysis by modeling multi-expert knowledge via fuzzy-logic rules, incorporating collected sensor data via neural networks, and combining domain knowledge and observed information via a novel hybrid reasoning approach.

The objective of this research into data visualization techniques was to model, represent, and display complex multi-dimensional terrestrial and atmospheric data and processes. Deception and confusion have been long-standing concerns in the discussion of computer-generated graphics [3]. Our objective was to create a presentation in which the form and style was consistent with the nature of the data. This environment allows multiple collaborative users to be fully immersed within a digital representation of integrated information. It has been seen that the user interaction components of immersive environment technologies are often poorly designed. As a result, many visually compelling environments are difficult to use and thus unproductive [4]. Our objective was to design and implement an environment that enables the user to easily interact with the earth science data, to facilitate the user’s innate human abilities to work collaboratively and assimilate multi-sensory information.

The following topics have been pursued:

- 1) Application of multi-dimensional, immersive visualization and interaction techniques to integrating earth science data products and the results of automated reasoning for improved interpretation and data discovery.
- 2) Development of visualization techniques to augment the presentation of ES geo-spatial, parametric, and temporal attributes within an immersive visualization environment.

Figure 1 presents the project concept of multi-source information integration and fusion for attaining improved contextual interpretation and understanding.

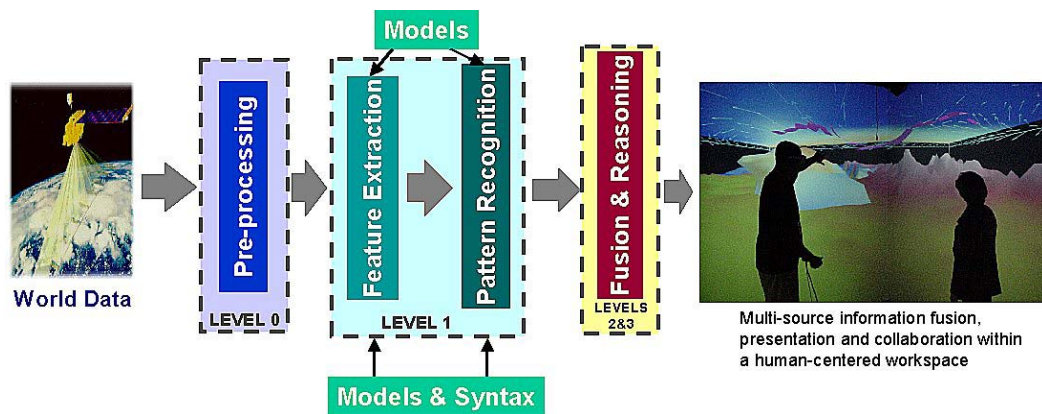


Figure 1: Multi-Source Earth Science Fusion and Visualization Objectives.

SUMMARY OF RESULTS

The research performed on this project demonstrated the potential for the successful application of hybrid reasoning approaches to the problems of earth science information processing, analysis and understanding. Applying hybrid reasoning techniques to the problem of predicting fire danger showed the capability to fuse non-commensurate information and domain knowledge to model, represent, and display complex multi-dimensional terrestrial and atmospheric data. While this effort was not uniquely focused as a solution to the problem of predicting fire danger, it presents the framework of a methodology that could be used for the solution of this and other earth science issues.

The immersive visualization provided an effective environment for users to experience the earth science information and data fusion results, interacting with the environment by means of well-practiced techniques and perceptual skills. By presenting the multi-source data in a geo-spatial, geo-temporal context in which the users could interactively navigate and manipulate the data, this environment showed potential for a variety of earth science analysis tasks. Although there are some limitations due to the current state of the technology, this demonstration showed that usable complex environments are possible. As data collection quantity and resolution continue to increase, the data volume will always exceed the technical capabilities so the community must continue to develop innovative and effective multi-sensory analysis techniques.

We believe that automated fusion methods supporting human analysis, such as those demonstrated by this project, will provide aids to reduce cognitive biases, improve the understanding of heterogeneous, multi-source data, and provide increased opportunities for data discovery.

REFERENCES

- [1] "Intelligent Systems (IS) Program Plan", NASA Ames Research Center Report, November, 2000.
- [2] M.J.M. Hall, S.A. Hall, and T. Tate, "Removing the HCI Bottleneck: How Human Computer Interfaces (HCI) Affects the Performance of Data Fusion Systems," Proceedings of the 2000 *National Symposium on Sensor Data Fusion (NSSDF)*, San Antonio, TX, June 2000.
- [3] A.J. David, "Presenting Visual Information Responsibly: Bad Graphs, Good Lessons," *ACM SIGGRAPH*, Vol. 33, No. 3, August 1999.
- [4] J.L. Gabbard, D. Hix, J.E. Swan II, "User-Centered Design and Evaluation of Virtual Environments," *IEEE Computer Graphics and Applications*, Vol. 19, No. 6, November/December, 1999, pages 51-59.

SYMPOSIA DISCUSSION – PAPER NO: 2

Author Name: Dr. David L. Hall, Pennsylvania State Univ., USA

Question:

How do you validate a hybrid approach? How do you know when you converge on a solution?

Author Response:

Develop the best set of rules possible, and then train a neural network to faithfully represent those rules. Then it is a matter of allowing the neural network to evolve with more data sets.

Question:

It is often difficult to extract a good percentage of the expert's knowledge, what are some ways of representing human intuition?

Author Response:

It is an iterative process. When you first apply the rules developed with an expert to data and get results, it might point to changes to the rules. The results of the evolved rules are taken back to the expert to confirm. Sometimes the expert did not think to include specific criteria, for example, or doesn't realize they were using a rule until it is pointed out.

Question:

The virtual reality cave method is very nice to look at, but appears awkward to do work in. You can't carry it around, where do you do your analysis, where is your workspace?

Author Response:

Different instantiations of the technology are suited to different users and applications. Oil exploration and automobile design are two of the industries utilizing this technology right now. It gives them the opportunity to bring people, possibly from different areas of expertise into the same environment and understanding. The choice of the display or instantiation of the technology utilized for a specific application should be one of the last things you do, after the requirements and application have been clearly defined.

Examples of uses cited:

- Mineral exploration where you can follow the path of the mine
- Medical applications such as the ability to remotely operate

PENNSTATE



Applied Research Laboratory

Multi-Sensory, Multi-Modal Concepts for Information Understanding

David Hall

School of Information Science and Technology

Timothy S. Shaw,

Applied Research Laboratory

Applied Research Laboratory
The Pennsylvania State University
P.O. Box 30
State College, PA 16804-0030

This research was funded in part by NASA Ames Research Center
under NASA Grant NCC 2-1211

Outline

- **The modern dilemma of knowledge acquisition**
- **A vision for information access and understanding**
- **Emerging concepts for data acquisition and understanding**
- **An experiment in data interpretation**
- **Research challenges**

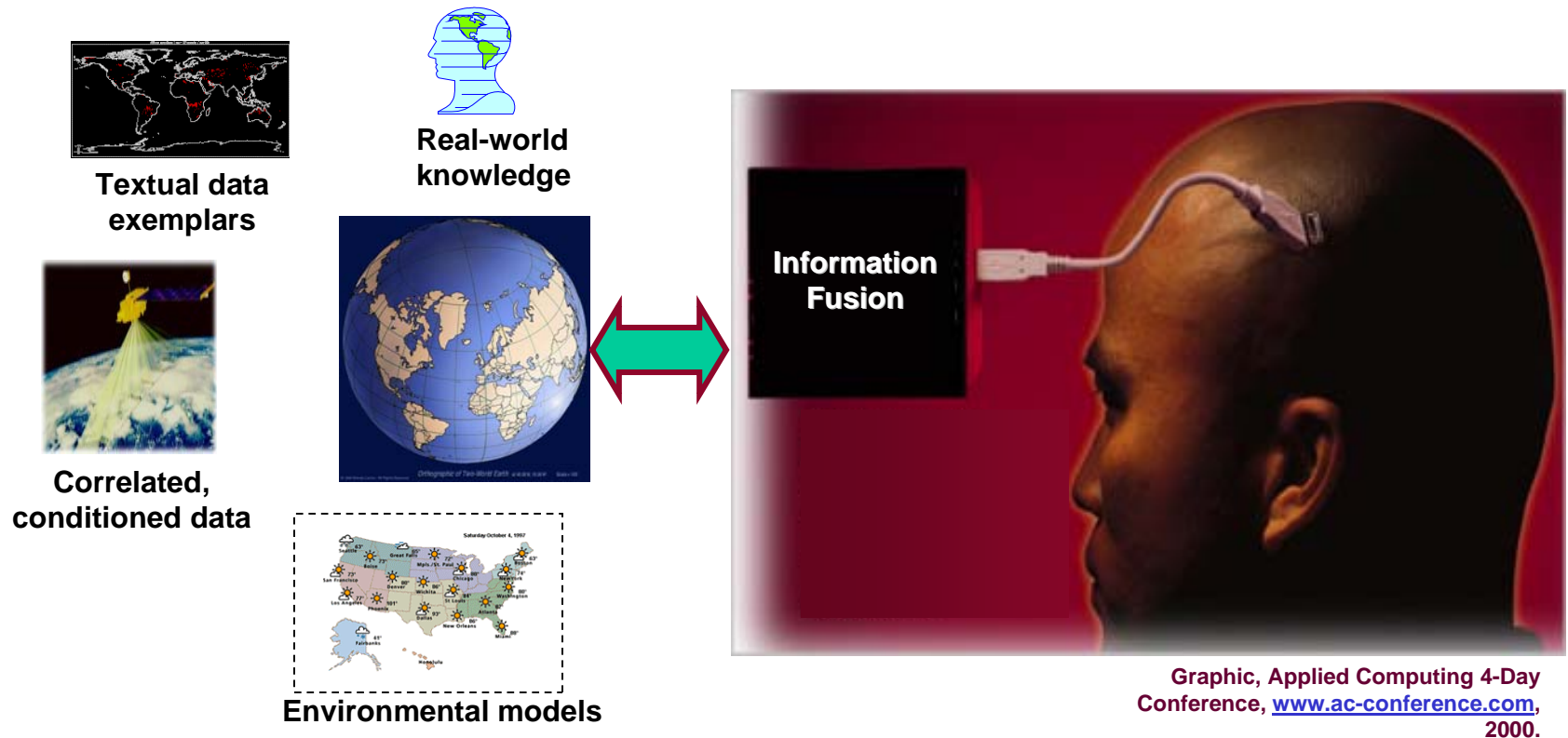
The modern dilemma of knowledge acquisition

- Ubiquitous sensing
 - Micro and nano-scale sensors
- Ultra-wide bandwidth communications
 - Wireless communications
 - Blurring of LANs and WANs
- Distributed computing and modeling
 - Internet accessible models and services (e.g., accu-weather)
- Overwhelming data rich (and model poor) environments
- Excess computing capability
- Cognitive versus computational barriers



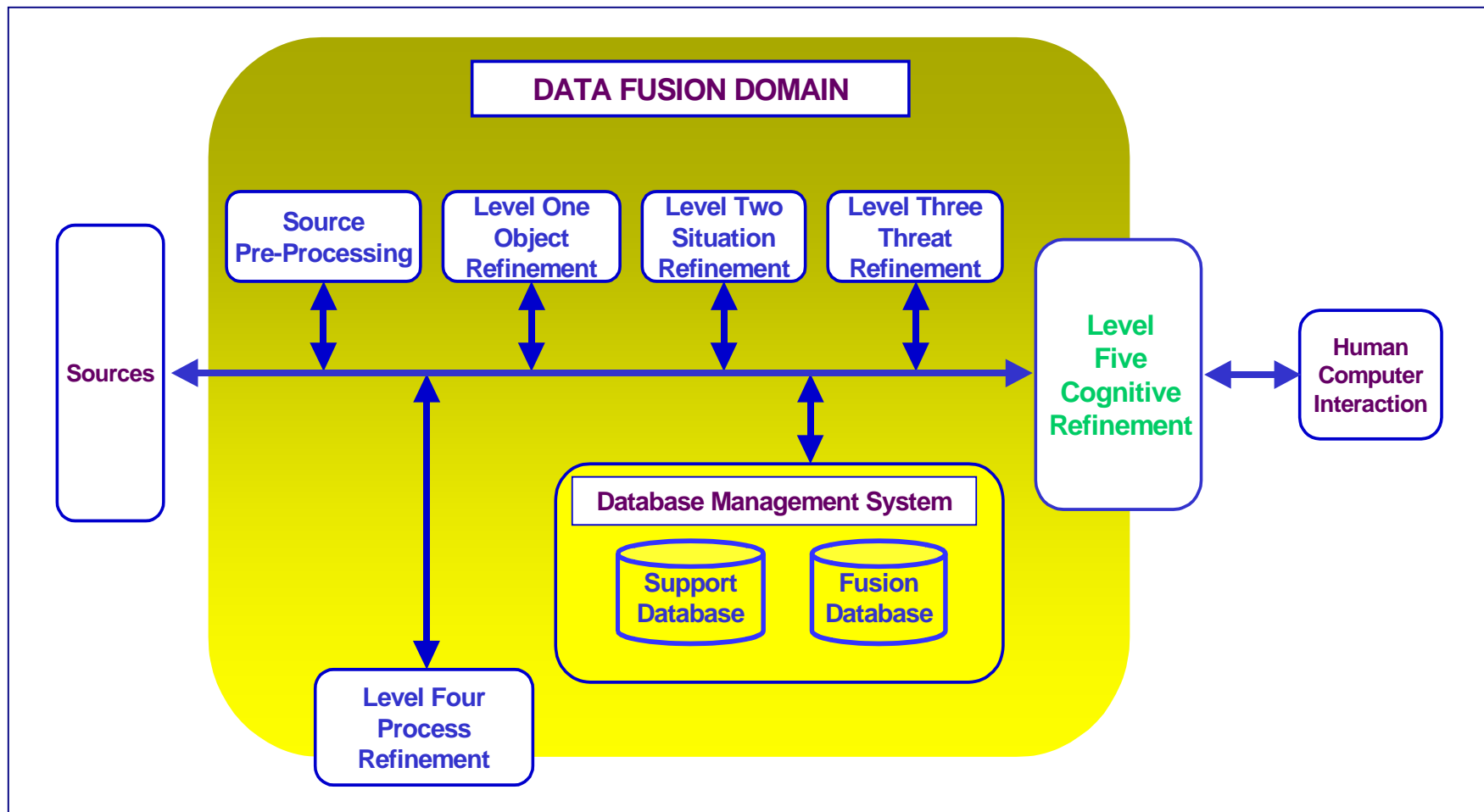
Why are we starved for knowledge while immersed in a sea of data?

Our Vision: *Bring Entire Earth's Resources to Individual Analyst*

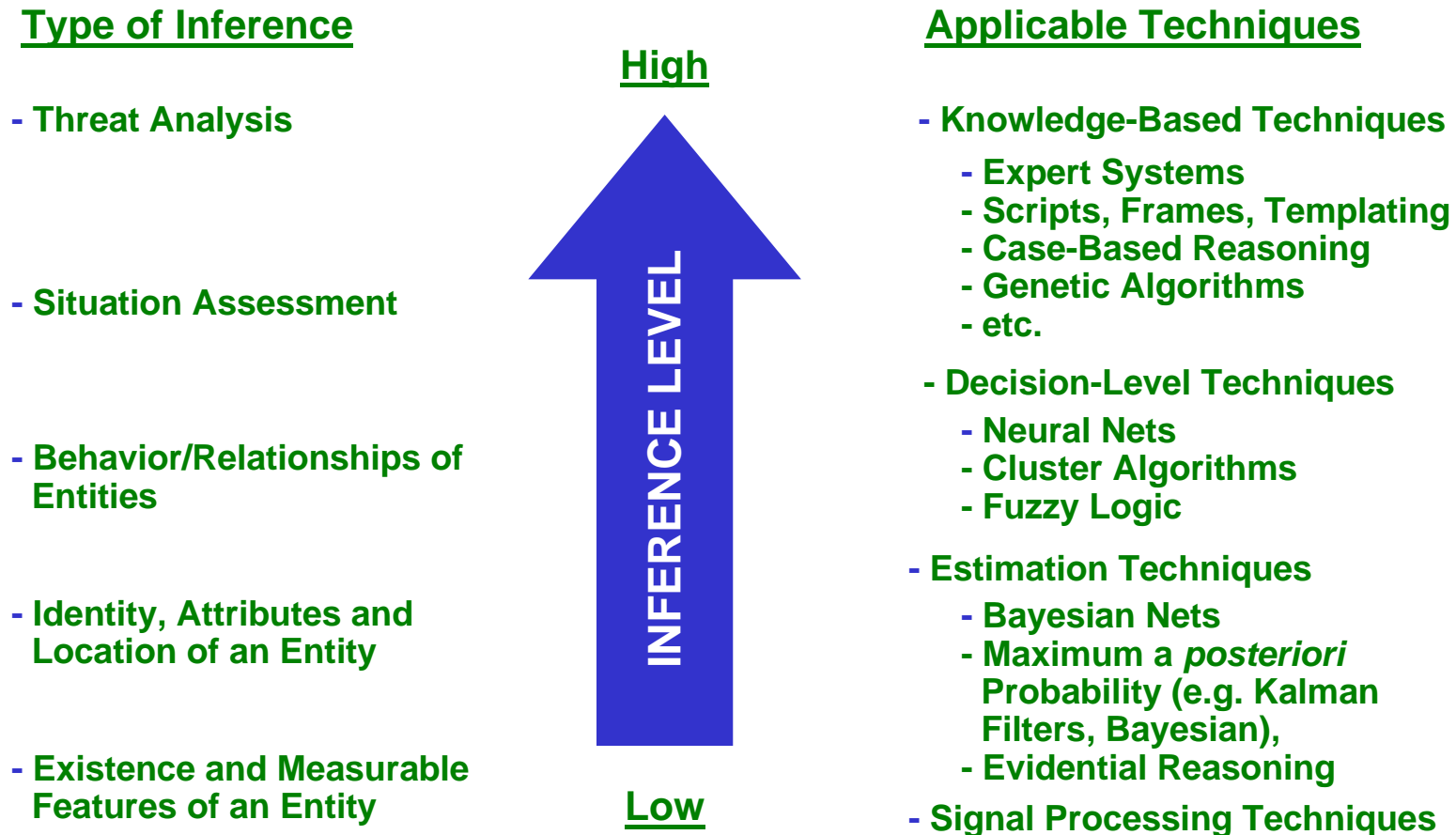


- Enabling technologies exist
- System effectiveness is affected by human characteristics
- Additional cognitive-centered research needed

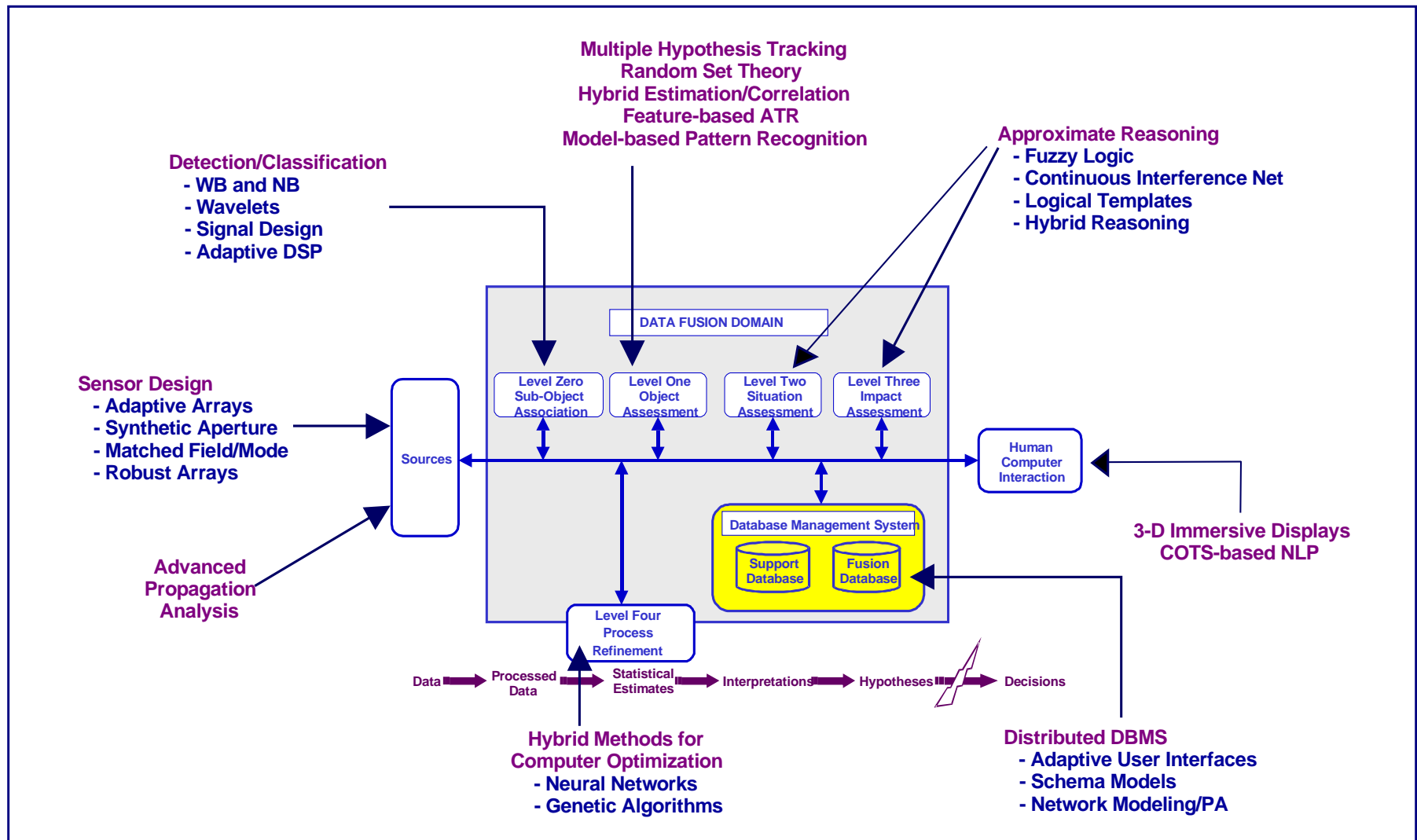
The JDL Data Fusion Model



Hierarchy of Inference Techniques



Trends in Data Fusion



An Experiment in Understanding Earth Science Data

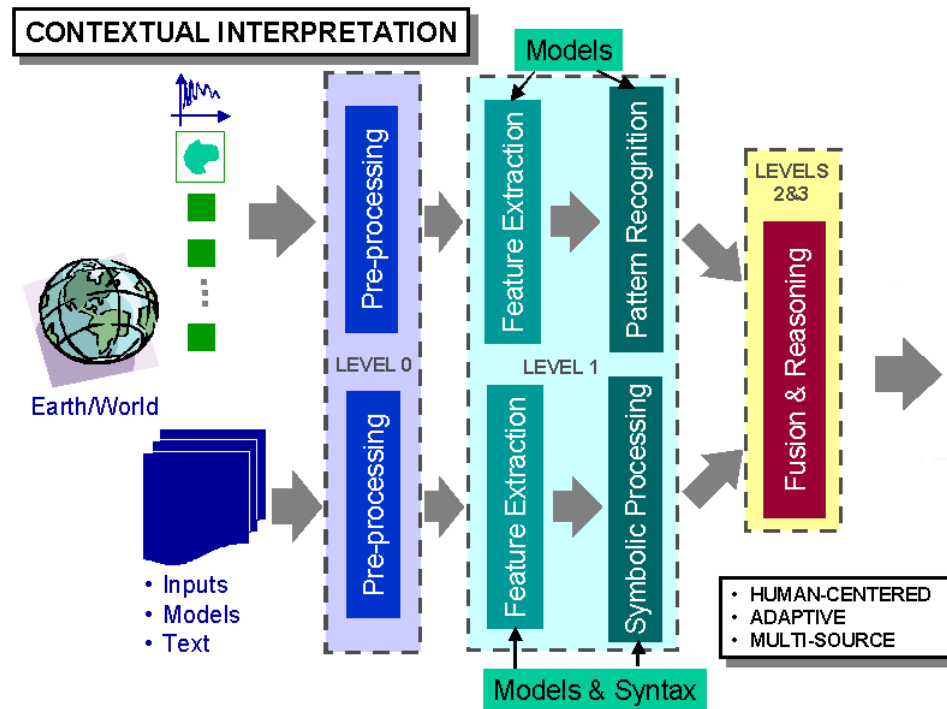
Contextual interpretation and human-center interaction with multi-source data

Innovations

- Combining sensor and textual data
- Contextual interpretation and understanding
- Human-centered presentation and user-profiling
- Transformation of image to semantic information

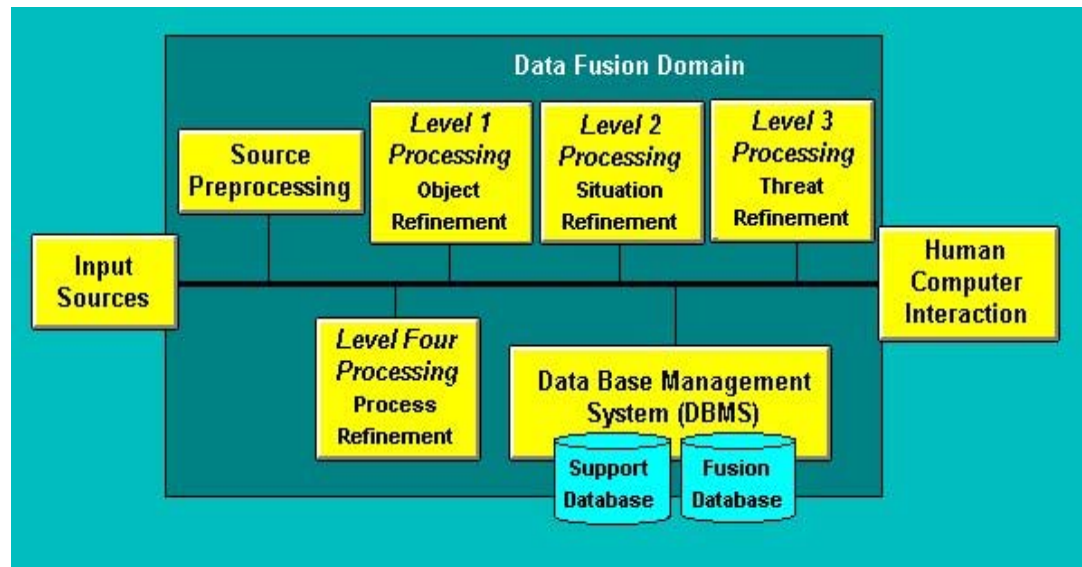
Techniques

- Immersive visualization environments
- Hybrid reasoning models
- Multi-modal human-computer interface
- Hierarchical wavelet-based image decomposition



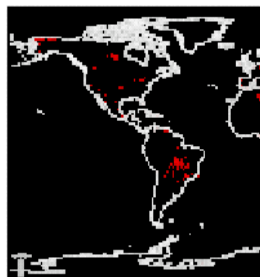
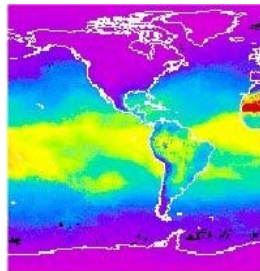
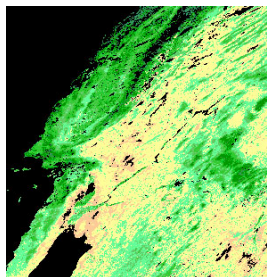
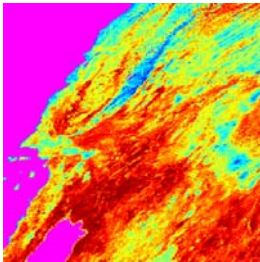
Multi-source information fusion and presentation enabling collaboration within a human-centered workspace

Key Research Strategies for Improved Data Understanding



- Fuse non-commensurate multi-source sensor data and information
- Perform automated reasoning to assist human analysis
- Explicitly design for human-in-the-loop data interpretation
- Utilize predictive analysis and alternative hypotheses evaluation

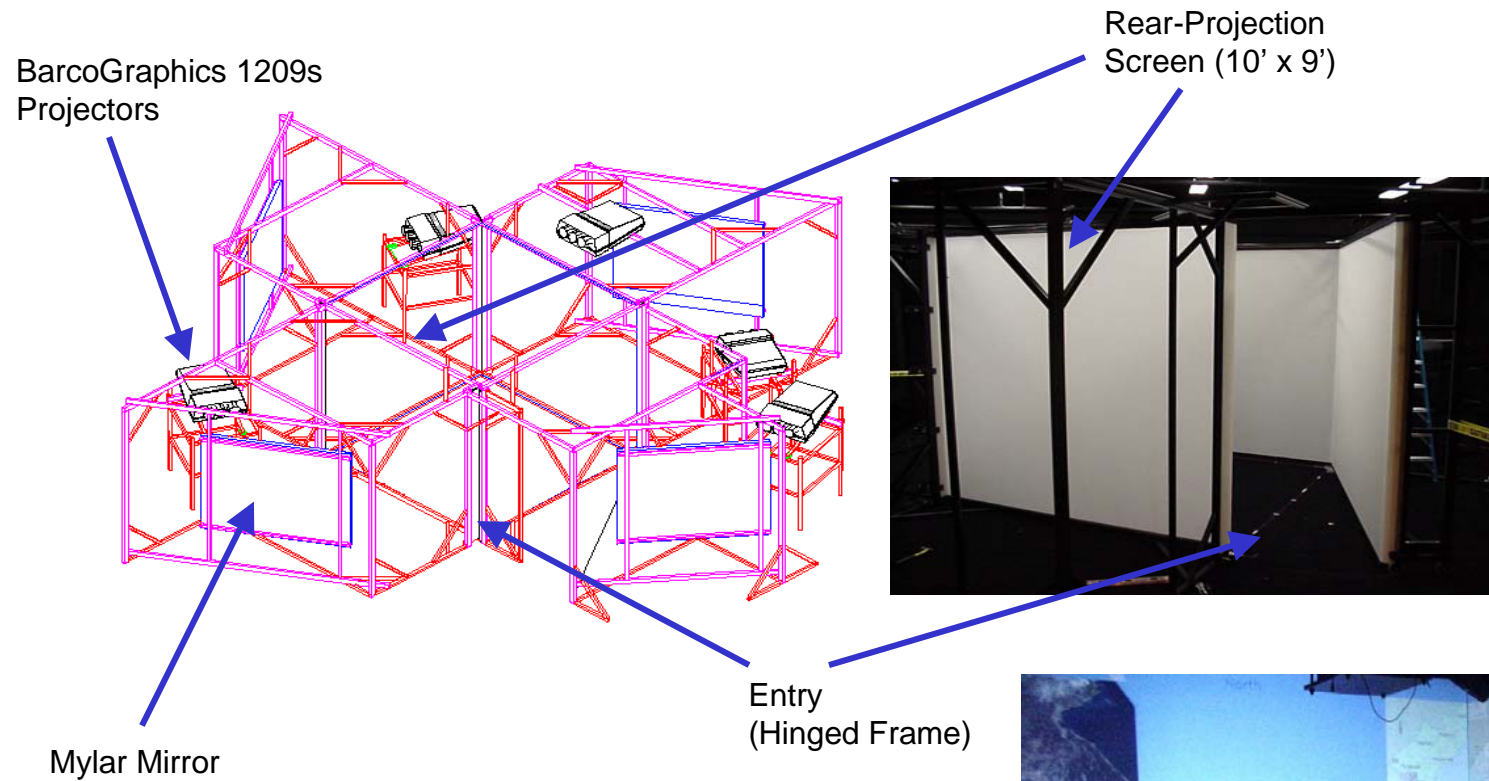
Remote-sensed Earth Science Data Sample



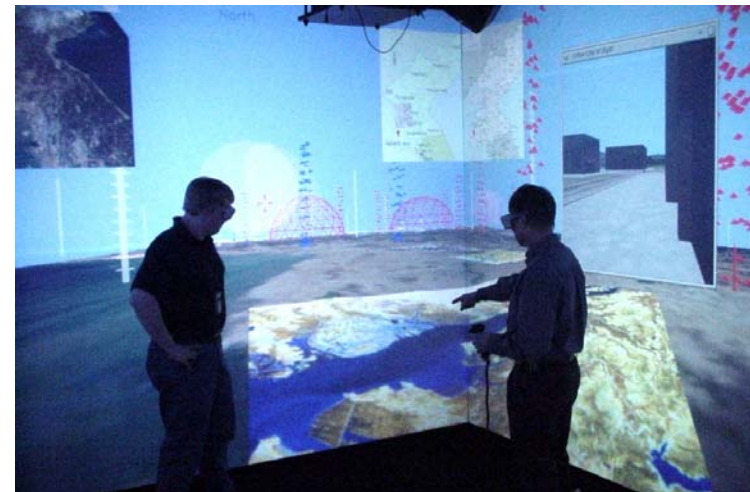
Data Product	Source	Data Format
<u>Land Surface Temperature</u> Day/night land temperature per grid	Terra satellite, MODIS sensor Bands 20, 22, 23, 29, 31, 32, 33	HDF-EOS, Integerized sinusoidal projection
<u>Leaf Area Index</u> One-sided leaf area per unit ground area	Terra satellite, MODIS sensor Bands 1 - 7	HDF-EOS, Integerized sinusoidal projection
<u>Precipitable Water</u> Column water vapor amounts	Terra satellite, MODIS sensor Bands 1, 2, 17, 18, 19	HDF-EOS, Equal angle grid
<u>Fire Event</u> Detected fire indication with time and location	ERS-1&2, ATSR sensor Bands of 1.6, 3.7, 11.0, 12.0 micrometers	Text report, Point location

Widely-available multi-source remote-sensed data and textual information can be fused to make interpretations and inferences using hybrid reasoning techniques.

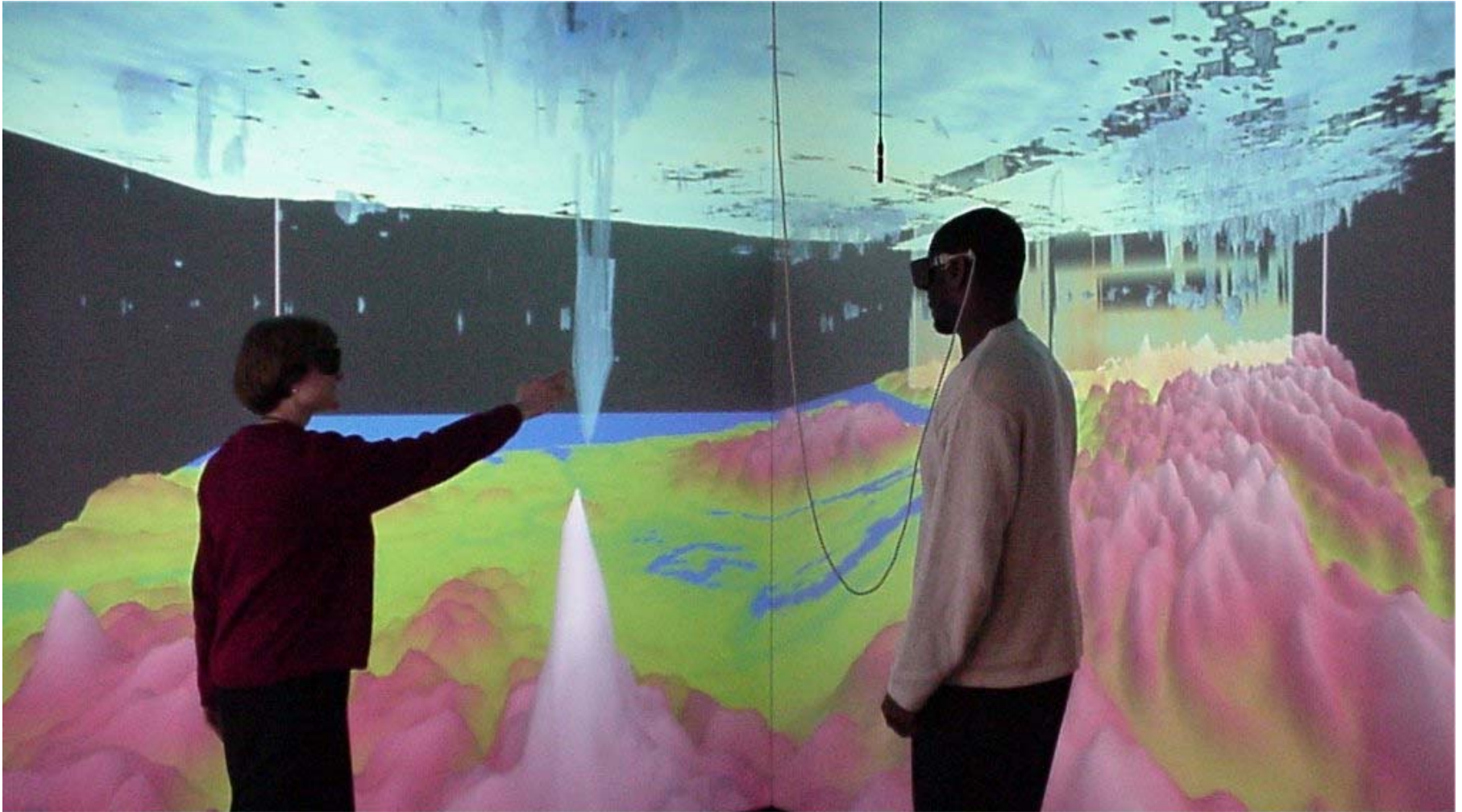
3-D Full Immersion Visualization Environment



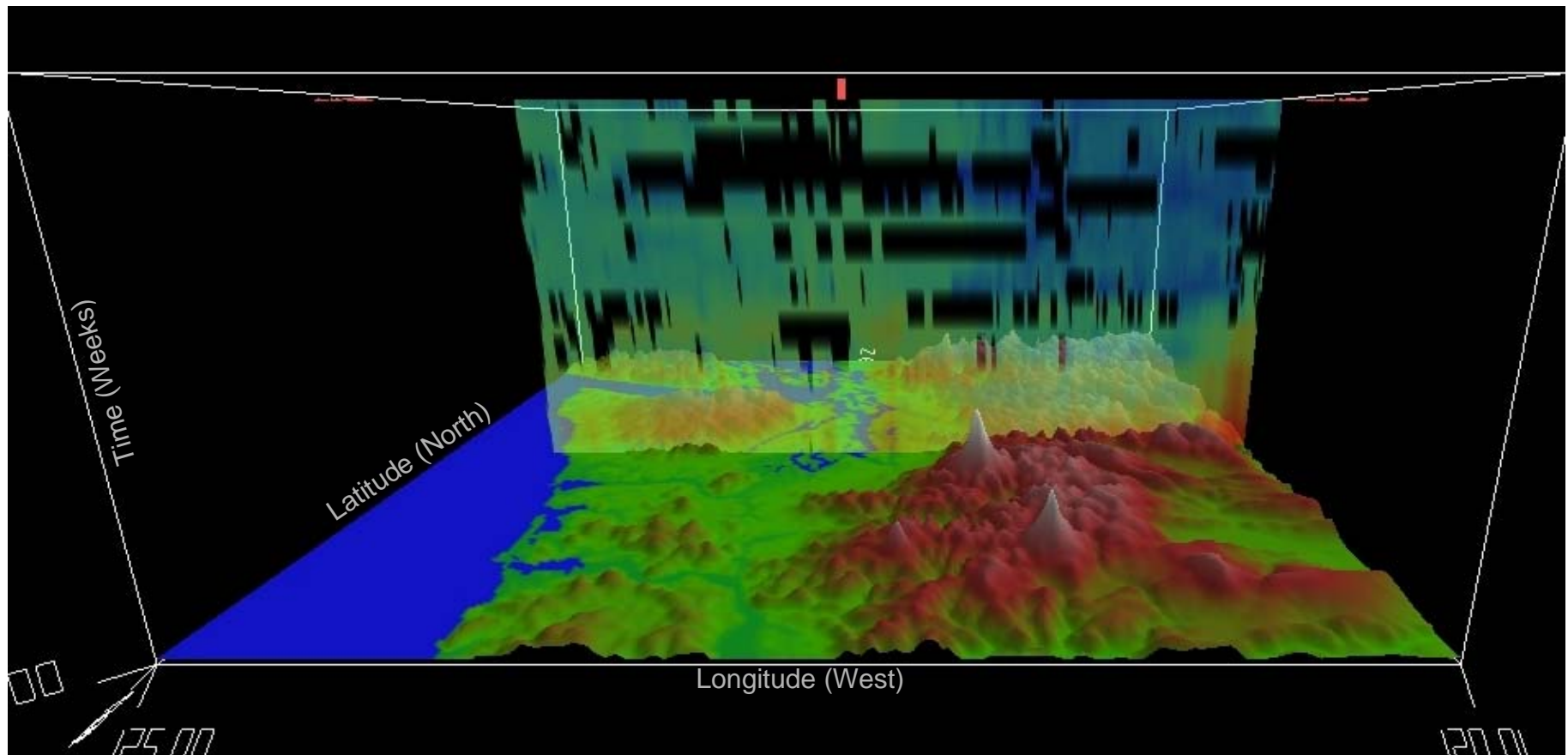
- Structure: 30' x 27' x 13'
- Center IPD Room: 10' x 10' x 9'
 - 4 walls, 360-degree surround images



ES Data Geo-temporal Visualization Example



ES Data Geo-temporal Visualization Example



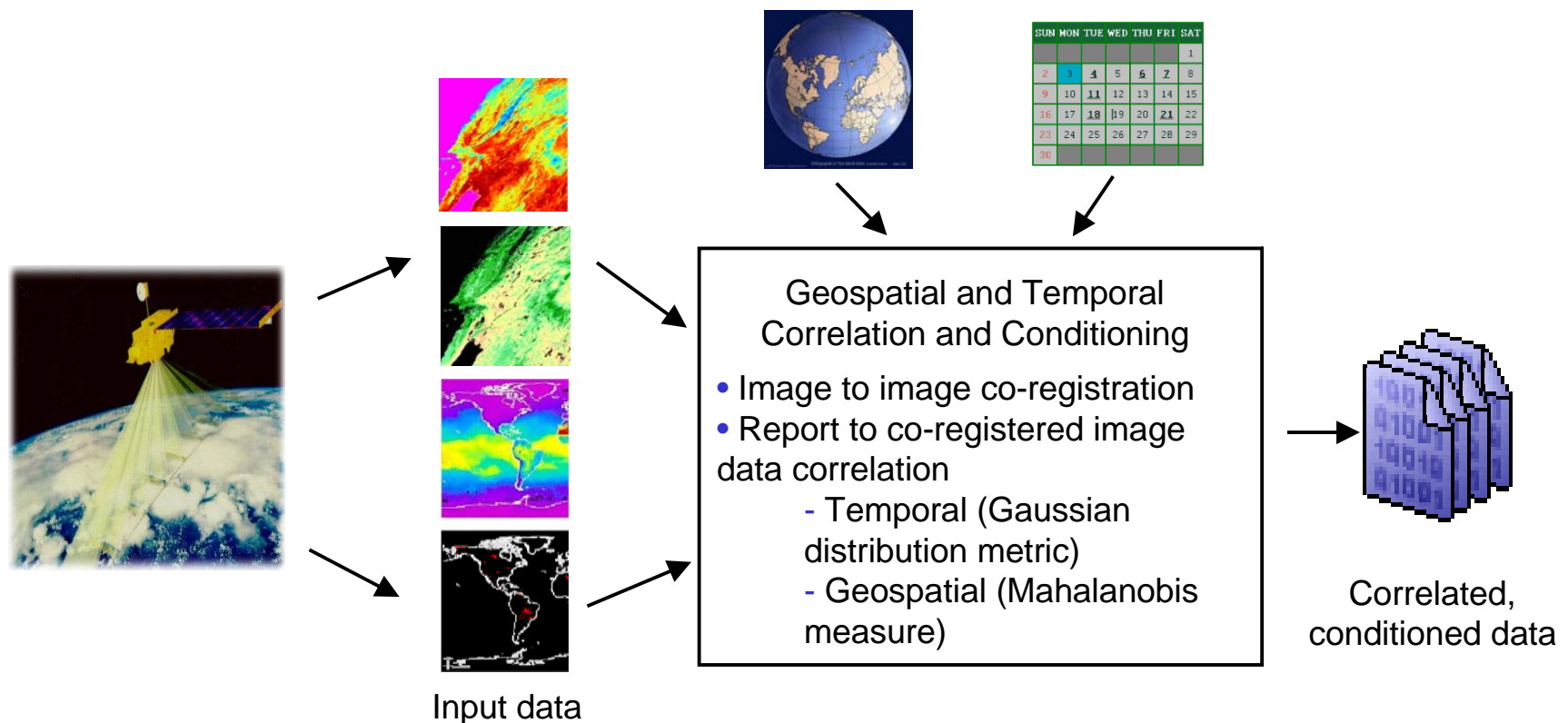
Data Preprocessing/Correlation (Level 0, 1)

Pre-processing

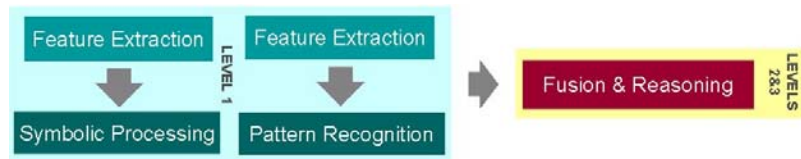
LEVEL 0

Pre-processing

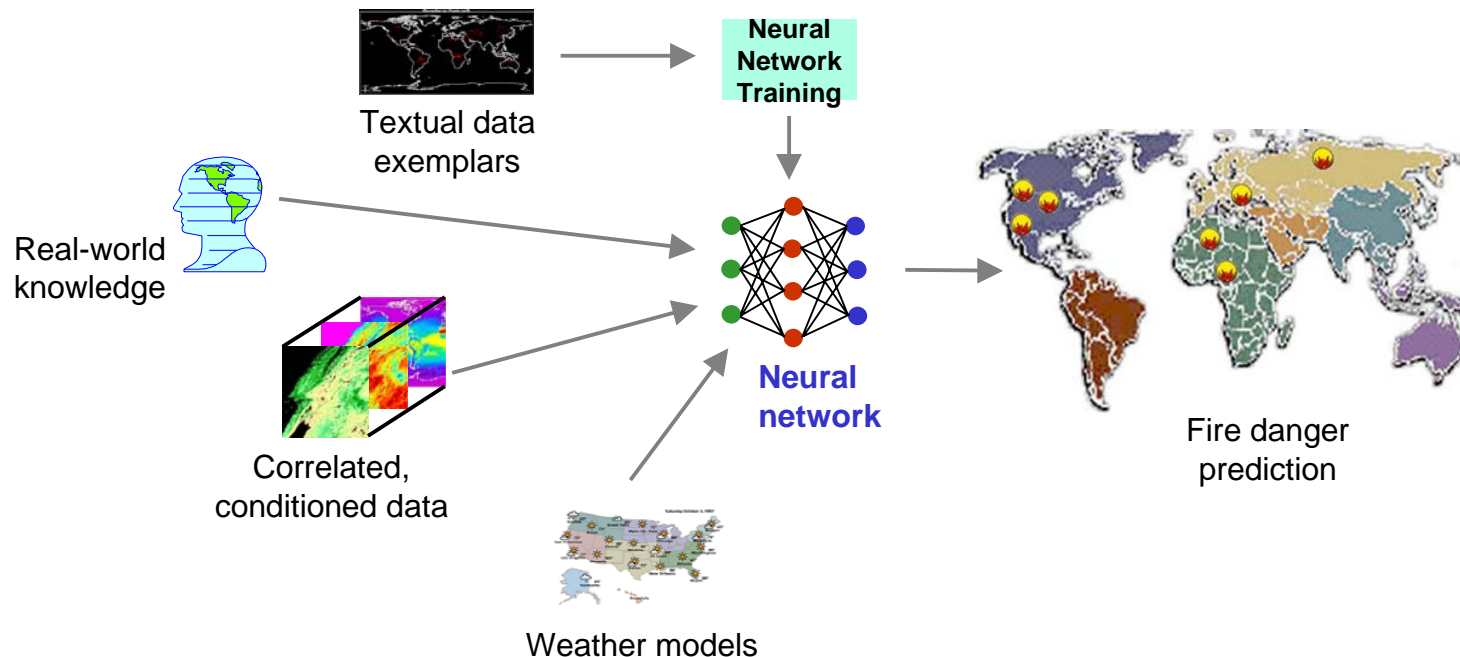
- **Data sets** - multiple types of Earth Science (ES) data
- **Data transformation** - ES data from HDF-EOS to binary
- **Level 1 correlation** - correlate the data values geo-spatially and temporally
- **Tools** - EOS Data Gateway, MODIS Re-projection Tool, HDF-EOS libraries



Symbolic Hybrid Reasoning (Level 2,3)



- **Hybrid reasoning** - combines explicit expert knowledge (from experts and models) with implicit knowledge “learned” from data
- **Hybrid knowledge representation** - overcomes limitations of individual techniques:
 - Limited data to train neural networks
 - Lack of specificity of explicit systems (e.g., rule-based systems)
- **Hybrid system performance** - more robust/less fragile than non-hybrid methods



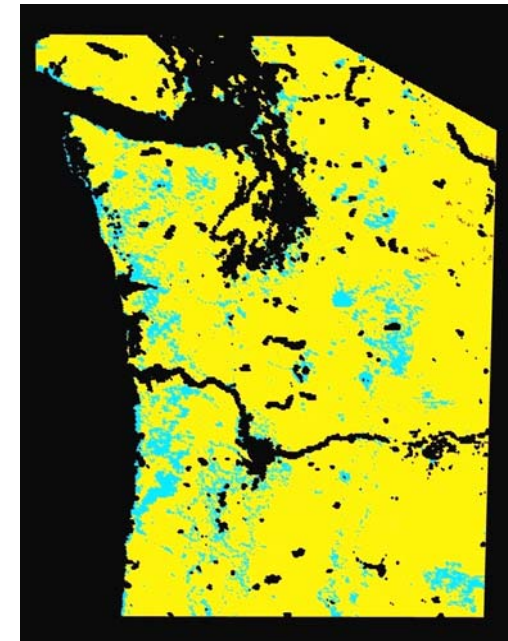
Explicit Reasoning

Explicit Rule Set

- **Multi-valued logic** - rules developed with 3 observable parameters that have an impact of fire danger
- **Input and output parameters** – data allowed to take values of Low, Medium, or High
- **Rule set** - 27 rules developed

Neural Network Training

- **Neural network**- network with 3 input neurons, 4 hidden neurons, and 1 output neuron
- **Network training** - trained using the back-propagation training method in MATLAB® Neural Network toolbox
- **Network performance** - performed 100% correct prediction on the rule base



Explicitly trained neural net of predicted fire danger over Washington State, Sept. 21, 2001

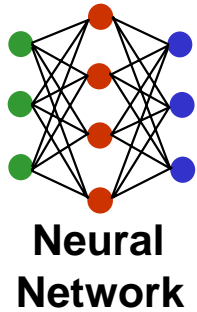
Fire danger:

Low - **blue**

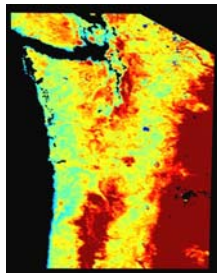
Medium - **yellow**

High - **red**

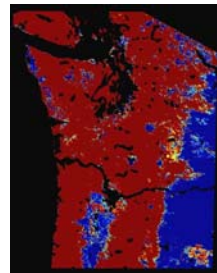
Hybrid Reasoning Results



- **Data set** - 17 observed fires in the region over the time period from March through November 2001
- **Explicit knowledge** - tested explicit neural net with half of the observed fire data; performed 20% correct prediction
- **Implicit knowledge** - implicit net trained on observed fire data, performed 80% correct prediction on test fire data
- **Hybrid results** - performed 87% correct prediction on test fire data with the hybrid neural net



LST

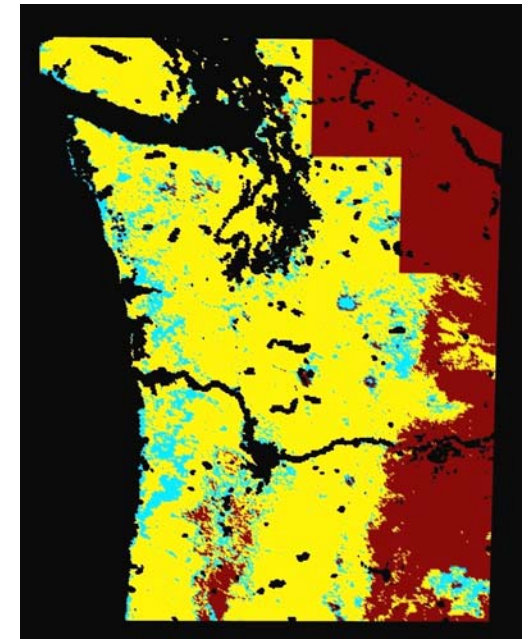


LAI



PWV
(lower resolution)

Observable parameters



Hybrid neural net of predicted fire danger over Washington State, September 21, 2001

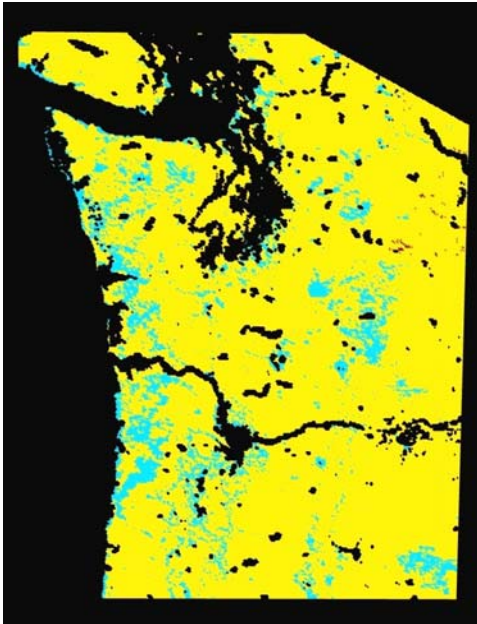
Fire danger:

Low - **blue**

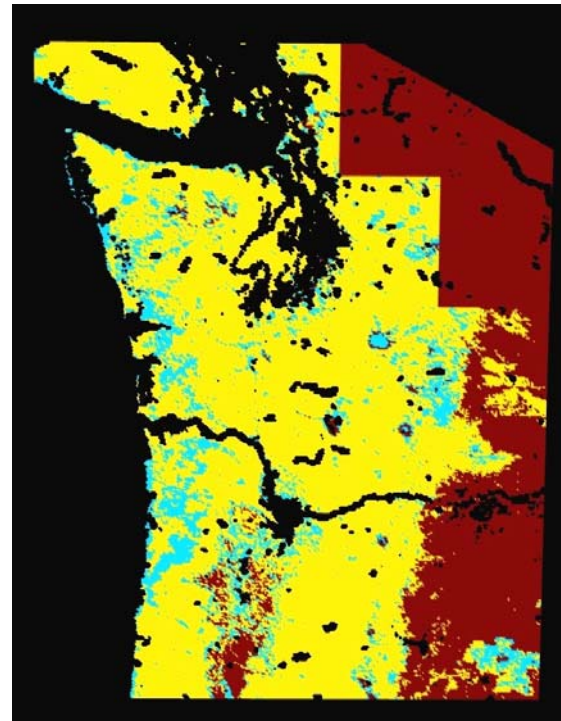
Medium - **yellow**

High - **red**

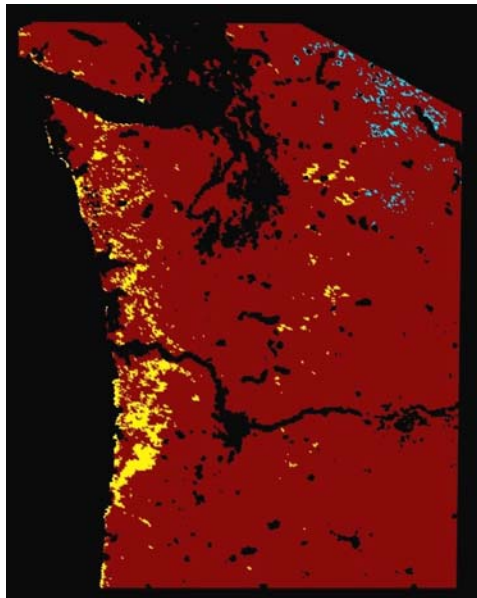
Comparison of Reasoning Performance



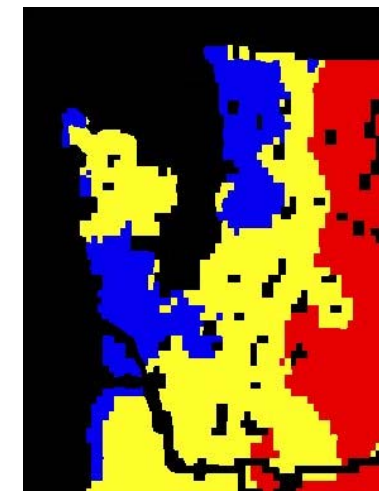
Explicit
 27% correct prediction



Hybrid
 87% correct prediction



Implicit
 80% correct prediction



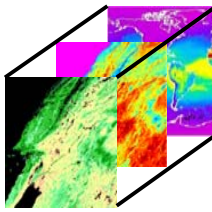
WFAS
 Observed Fire Danger

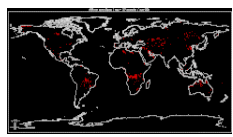
Fire danger:
 Low - **blue**
 Medium - **yellow**
 High - **red**

Level 3 Prediction Concept

Fusion & Reasoning

LEVELS
2&3


Correlated, conditioned data
Georegistered
Time-stamped


Textual reports
Geographic location
Time-stamped

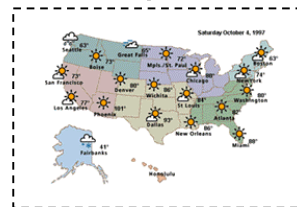
Data Fusion and Interpretation

- Feature-level and report-level fusion using hybrid reasoning system
- Explicit real-world knowledge representation (via fuzzy logic rules)
- Incorporation of implicit information using real data exemplars

Predictive and Causal Modeling

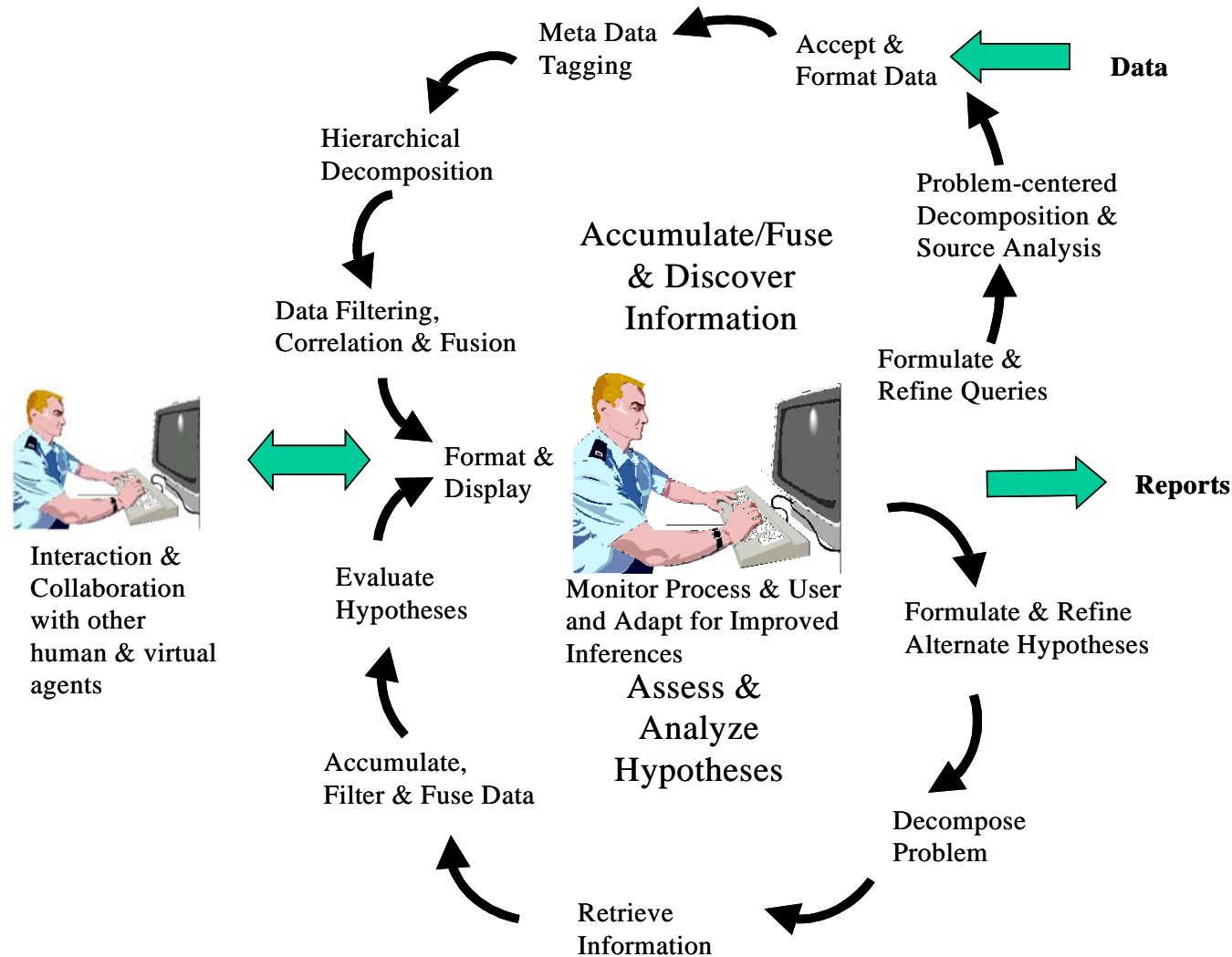
- Explicit causal hypotheses (via fuzzy logic rules)
- Incorporation of implicit data by evolving the neural network
- Multi-time scale recurrent neural networks incorporating predictive meteorological models

- Interpretation
- Predictions



Weather models

The Balance of Knowledge Discovery and Analysis



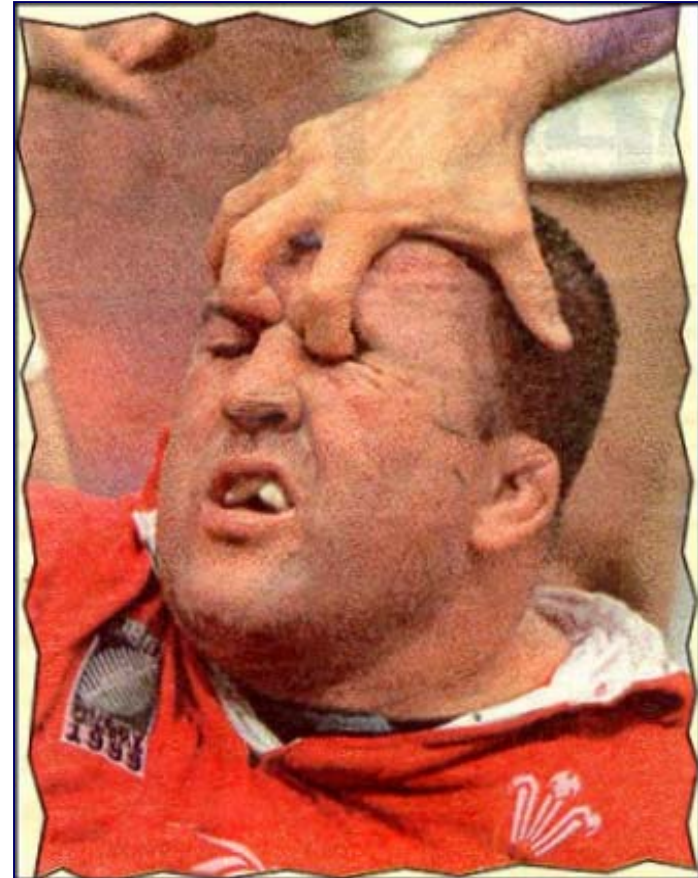
Research Issues and Challenges

- Geo-spatial and temporal resolution of the data
- Availability of training data
- Multi-times scale and asynchronous data
- Prediction intervals and predictability horizons
- Incorporation of multi-expert knowledge
- “Brittleness” of prediction and reasoning
- Incorporation of negative reasoning
- Default reasoning
- Approach for indeterminate & unavailable information
- Human-in-the-loop processing and multi-person collaboration
- Development of cognitive aids for interpretation
- Multi-sensory representations of uncertainty

This problem provides a rich source of continuing challenges across multiple disciplines.

Summary

- With some imagination and near-term innovations, information fusion and understanding may not be so difficult



Picture: Stuart MacFarlane,
London, Daily News, October 1999.